Goal: Given a set S, determine the cardinality $\#S$.

Important Terms

We let $[n] \coloneqq \{1, 2, \cdots, n\}$; observe that for all $n \in \mathbb{N}_0$ we have $\#[n] = n$.

A *partition* of a set S is a collection $\{B_1, B_2, \cdots, B_r\}$ of sets satisfying the following conditions.

1. For all $k \in [r]$ we have $B_k \neq \emptyset$.
2. The union $\displaystyle\bigcup_{k=1}^{r} B_k = S$.
3. For all $i, j \in [r]$, if $B_i \cap B_j \neq \emptyset$, then $i = j$.[1]

The elements B_k of a partition are called the *blocks* of the partition. Thus $\{\{1, 2, 3\}, \{4, 7\}, \{5, 6\}\}$ is a partition of $[7]$ with blocks $\{1, 2, 3\}$, $\{4, 7\}$, and $\{5, 6\}$.

A *word* in an *alphabet* A is a sequence $a_1 a_2 \cdots a_n$ of symbols of the set A. The *length* of a word is the number of characters it uses (including repetitions). Thus the length of the word $MISSISSIPPI$ is 11. Note that the order of the letters in a word matters! Thus $MISSISSIPPI$ and $IPPISSISSIM$ are different words despite the fact that they use the same letters. Also, **a word need not be an English word in this context**.

Two words W_1 and W_2 are *anagrams* if one can be obtained from the other by rearranging their letters. Thus $MAGMA$ an $GAMMA$ are anagrams, but $MAGMA$ and $ANAGRAM$ are not anagrams.

A *palindrome* is a word equal to its reverse. Thus $AMASSAMA$ is a palindrome, but $MISSISSIPPI$ is not.

A *permutation* of a set S is a bijection $f : S \to S$.

Bijection Principle

The most basic way of counting a set is to directly construct a bijection between that and another set for which we already know the cardinality. We will call this the *Bijection Principle*.[2]

Proposition 1 (Bijection Principle)**.** *Let S and T be sets. We have $\#S = \#T$ if and only if there is a bijection $f : S \to T$.*

Example 1. For all $n \in \mathbb{N}_0$, the set $S_n \coloneqq \{k \in \mathbb{Z} : -n \leq k \leq -1\}$ has cardinality n. The functions below are inverses.

$$f : S_n \to [n] : k \mapsto -k \qquad\qquad g : [n] \to S_n : k \mapsto -k$$

Example 2. For all $n \in \mathbb{N}_0$, the set $T_n \coloneqq \{k \in \mathbb{Z} : -n \leq k \leq n\}$ has cardinality $\#T_n = 2n+1$. Construct a bijection.

Problem 1. Let A be a finite alphabet. Prove that the set of length $2n$ palindromes in A is in bijection with the set of length $2n - 1$ palindromes in A.

Addition Principle

One approach to counting a complicated set S is to break it into more manageable pieces, count those smaller pieces, and use those counts to count the full set. This idea is the *Addition Principle*, which we can state as follows.

Proposition 2 (Addition Principle)**.** *Let $A_1, A_2, \cdots, A_n$ be (finite) sets and suppose $A_i \cap A_j = \emptyset$ whenever $i \neq j$. Then the set $S = \displaystyle\bigcup_{k=1}^{n} A_k$ has cardinality*

$$\#S = \#\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \#A_k.$$

Example 3. Consider the set $S = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : x \in [3], y \in [5]\}$. Then we can split this set S into 3 peices, $A_k \coloneqq \{(x, y) \in S : x = k\}$ for $k \in \{1, 2, 3\}$. Moreover, because 1, 2, and 3 are all distinct, $A_i \cap A_j = \emptyset$ whenever $i \neq j$. Now $\#S = \#(A_1 \cup A_2 \cup A_3) = \#A_1 + \#A_2 + \#A_3$ by the Addition Principle. For each $k \in [3]$ the function $f_k : A_k \to [5] : (i, y) \mapsto y$ is bijective (why?). Hence $\#S = \#A_1 + \#A_2 + \#A_3 = 5 + 5 + 5$ by the Bijection Principle.

[1] Equivalently $i \neq j$ implies $B_i \cap B_j = \emptyset$.

[2] I state this as a proposition below, but you could alternatively take this as a definition of cardinality.

The Addition Principle is very powerful. To illustrate this power, we prove the following (a generalization of the example above).

Proposition 3. *For all (finite) sets S and T we have $\#(S \times T) = (\#S) \cdot (\#T)$.*

Proof. Let S and T be arbitrary finite sets; write $\#S = n$. Thus we can write $S = \{s_1, s_2, \cdots, s_n\}$, listing the elements of S. Now we define $A_k \coloneqq \{(s_k, t) : t \in T\}$ for all $k \in [n]$. Note that if $i \neq j$, then $s_i \neq s_j$, so $(s_i, t) \neq (s_j, t')$ for all $t, t' \in T$; thus $A_i \cap A_j = \emptyset$ whenever $i \neq j$. Moreover, for all $(s, t) \in S \times T$ there is an k such that $s = s_k$; thus $(s, t) = (s_k, t) \in A_k$. In particular we see $S \times T = \bigcup_{k=1}^{n} A_k$. Hence by the Addition Principle

$$\#(S \times T) = \#\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \#A_k.$$

Moreover, for all $k \in [n]$ there is a bijection $f_k : A_k \to T : (s_k, t) \mapsto t$ (**Exercise**: Prove that f_k is a bijection.). Hence by the Bijection Principle and basic arithmetic we complete the proof with the calculation

$$\#(S \times T) = \sum_{k=1}^{n} \#A_k = \sum_{k=1}^{n} \#T = \left(\sum_{k=1}^{n} 1\right) \#T = n \cdot \#T = (\#S) \cdot (\#T). \qquad \square$$

Corollary 1. *For all (finite) sets $S_1, S_2, \cdots, S_n$ we have*

$$\#(S_1 \times S_2 \times \cdots \times S_n) = \prod_{k=1}^{n} \#S_k.$$

Proof. Exercise. **Hint**: Use induction on n, the above proposition, and the Addition Principle. $\square$

Product Principle

Sometimes complicated sets can be naturally counted by making a sequence of choices to "build" each of the elements of S. Counting such sets is easy, provided the numbers of choices remain the same at every step. This idea is the *Product Principle.*

Proposition 4 (Product Principle)**.** *If every element of a (finite) set S can be constructed uniquely by a sequence of n choices such that at the k^{th} choice there are c_k options independent of all previous choices, then*

$$\#S = \prod_{k=1}^{n} c_k.$$

The idea here comes from the same type of logic we used to prove $\#(A \times B) = (\#A) \cdot (\#B)$; essentially, the product principle descrbes a bijection from the set S to the cartesian product of the set of choices made.

Example 4. Consider the set of words of length 4 using letters of the English alphabet. Each such word is built by first choosing the first letter (26 options), next choosing the second letter (26 options), then choosing the third letter (26 options), and finally choosing the last letter (26 options). As every word is uniquely determined by this sequence of choices, there are $26 \cdot 26 \cdot 26 \cdot 26 = 26^4$ words of length 4 using the letters of the Engllish alphabet.

Proposition 5. *Let S and T be (finite) sets. There are $(\#T)^{(\#S)}$ functions with domain S and codomain T.*

Proof. Let S and T be finite sets, and let $\#S = n$. Thus we can write $S = \{s_1, s_2, \cdots, s_n\}$ listing out the elements of S. Now we can express the set of all functions from S to T as a sequence of n choices. To build a function $f : S \to T$ first we choose a sequence $(t_1, t_2, \cdots, t_n)$ of elements of T and define $f(s_k) = t_k$ for all $k \in [n]$. Furthermore, this uniquely determines the function f. For each $k \in [n]$ there are $\#T$ options for t_k. Thus there are $(\#T)^n = (\#T)^{(\#S)}$ such sequences by the Product Principle, and hence the same number of functions of the desired type. $\square$

Proposition 6. *There are $n!$ permutations of $[n]$ for all $n \in \mathbb{N}_0$.*

Proof. Exercise. **Hint**: Use induction on n and the Product Principle. $\square$

Problem 2. Let S and T be finite sets. How many injections are there with domain S and codomain T?

Binomial Coefficients

Let $n, k \in \mathbb{N}_0$. The *binomial coefficient* is defined by

$$\binom{n}{k} := \#\left\{S \subseteq [n] : \#S = k\right\}.$$

Typically we read $\binom{n}{k}$ as "n choose k". Binomial coefficients have a lot of interesting properties.

Problem 3. Prove each of the following.

1. For all $n \in \mathbb{N}_0$ we have $\binom{n}{0} = 1$.
2. For all $n \in \mathbb{N}_0$ and all $k \in \{0\} \cup [n]$ we have $\binom{n}{k} = \binom{n}{n-k}$.
3. For all $n \in \mathbb{N}_0$ we have $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$.
4. For all $n \in \mathbb{N}_0$ and all $k \in \{0\} \cup [n]$ we have $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

Problem 4. Use binomial coefficients to count the number of anagrams of the word $MISSISSIPPI$.

Inclusion-Exclusion Principle

We've already seen that some counting problems are easiest too approach by breaking a set into pieces; the addition principle states that by partitioning a set and counting the blocks separately we can count the full set by adding the counts of the blocks. Sometimes it is not natural to build a partition of a set, but there might be a natural way to split the set into several parts and count those; however, we still need a way to reassemble the counts of these pieces to count the full set. Here enters the *Inclusion-Exclusion Principle*.

Proposition 7 (Inclusion-Exclusion Principle)**.** *For all finite sets A and B we have $\#(A \cup B) = \#A + \#B - \#(A \cap B)$.*

The inclusion-exclusion formula remedies of overcounting by subtracting off the twice-counted members.

Example 5. A class of 300 freshmen has 100 maths majors and 150 computer science majors. If 40 students are maths and computer science dual majors, then there are $100 + 150 - 40 = 210$ freshmen in maths or computer science.

There is a nice general inclusion-exclusion formula for bigger collections of finite sets; this we state below.

Proposition 8 (General Inclusion-Exclusion Principle)**.** *For all finite sets $A_1, A_2, \cdots, A_n$ we have*

$$\#\bigcup_{k=1}^{n} A_k = \sum_{K \subseteq [n]} (-1)^{\#K} \# \bigcap_{k \in K} A_k.$$

The general inclusion-exclusion formula corrects overcounting first by subtracting off intersections, corrects the resulting undercounting by adding the triple intersections, and repeats until correct. For $n = 3$, the formula is

$$\#(A \cup B \cup C) = \#A + \#B + \#C - \#(A \cap B) - \#(A \cap C) - \#(B \cap C) + \#(A \cap B \cap C).$$

Draw a Venn Diagram to convince yourself that this formula is counting correctly (or just prove it).

Example 6. Let S be the set of length 4 words in the alphabet $A = \{0\} \cup [9]$ having the subword 01. Define

$$\begin{aligned} S_1 &= \{w \in S : w = 01xy \text{ for some } x, y \in A\} \\ S_2 &= \{w \in S : w = x01y \text{ for some } x, y \in A\} \\ S_3 &= \{w \in S : w = xy01 \text{ for some } x, y \in A\} \end{aligned}$$

Now notice that $\#S_1 = \#S_2 = \#S_3 = 10^2$ as we have 10 choices for each of x and y. Now $S_1 \cap S_2 = \emptyset = S_2 \cap S_3$ by comparing the requirements for w; thus also $S_1 \cap S_2 \cap S_3 = \emptyset$. Finally, $S_1 \cap S_3 = \{0101\}$; thus

$$\begin{aligned} \#S &= \#S_1 + \#S_2 + \#S_3 - \#(S_1 \cap S_2) - \#(S_1 \cap S_3) - \#(S_2 \cap S_3) + \#(S_1 \cap S_2 \cap S_3) \\ &= 100 + 100 + 100 - 0 - 1 - 0 + 0 = 299. \end{aligned}$$

Example 7. Let S be the set of length 5 words in the alphabet $A = \{0\} \cup [9]$ with a subword either 01, 70, or 37. Define the following.

$$\begin{aligned} X &= \{w \in S : 01 \text{ is a subword of } w\} \\ Y &= \{w \in S : 70 \text{ is a subword of } w\} \\ Z &= \{w \in S : 37 \text{ is a subword of } w\} \end{aligned}$$

First we count the sets X, Y, and Z. We will count the set $S(a,b) = \{w \in S : ab \text{ is a subword of } w\}$ where $a \neq b$. To count $S(a,b)$, we define four new sets.

$$\begin{aligned} S_1 &= \{w \in S(a,b) : w = abxyz \text{ for some } x,y,z \in A\} \\ S_2 &= \{w \in S(a,b) : w = xabyz \text{ for some } x,y,z \in A\} \\ S_3 &= \{w \in S(a,b) : w = xyabz \text{ for some } x,y,z \in A\} \\ S_4 &= \{w \in S(a,b) : w = xyzab \text{ for some } x,y,z \in A\} \end{aligned}$$

Primarily notice that $\#S_i = 10^3$ for all $i \in [4]$. Moreover, $S_1 \cap S_2 = S_2 \cap S_3 = S_3 \cap S_4 = \emptyset$. On the other hand, $\#(S_1 \cap S_3) = \#(S_1 \cap S_4) = \#(S_2 \cap S_4) = 10$. Finally, $S_i \cap S_j \cap S_k = \emptyset$ for all distinct $i,j,k \in [4]$. Thus we have

$$\#S(a,b) = 10^3 + 10^3 + 10^3 + 10^3 - 0 - 10 - 10 - 0 - 10 - 0 - 0 = 3070$$

Now note that $X = S(0,1)$, $Y = S(7,0)$, and $Z = S(3,7)$. Now we compute the sizes of the intersections of X, Y, and Z; again, we will be tricky and compute this in terms of the more general sets $S(a,b)$.

Suppose a, b, and c are all distinct. Note that $S(a,b) \cap S(b,c)$ has a natural breakdown into the sets

$$M = \{w \in S : abc \text{ is a subword of } w\} \qquad \text{and} \qquad N = \{w \in S : abc \text{ is not a subword of } w\}.$$

An word in S is constructed by choosing a position for the a in the abc string (3 options) and choosing the remaining two characters (10^2 options); thus $\#M = 300$. Every word in N is one of the following for some $x \in A$:

$$abbcx \qquad abxbc \qquad xabbc \qquad bcabx \qquad bcxab \qquad xbcab$$

Notice that if xbc is a subword of w, then $x \neq a$; similarly, if abx is a subword of w, then $x \neq c$. Thus we have 10, 9, 10, 9, 10, and 9 such words in the cases above. Hence $\#N = 57$, which yields $\#(S(a,b) \cap S(b,c)) = \#M + \#N = 357$.

Now suppose a,b,c,d are all distinct. Then every element of $S(a,b) \cap S(c,d)$ can be constructed from a triple (x,y,z) where $\{x,y,z\} = \{ab, cd, e\}$ with $e \in A$. We count such triples by first choosing an order to put ab and cd (2 options), choosing a position for e (3 options), and choosing the value of $e \in A$ (10 options). Hence $\#(S(a,b) \cap S(c,d)) = 60$.

Thus we have shown $\#(X \cap Y) = 357 = \#(Y \cap Z)$ and $\#(X \cap Z) = 60$.

Finally, to count $X \cap Y \cap Z$, notice that one of 701 or 370 is a subword of every member of $X \cap Y \cap Z$; otherwise there are too many characters in the word. Let P denote the set of words in S with 701 as a subword, and let Q denote the set of words with 370 as a subword. Now $\#P = 300 = \#Q$ as we counted above (when computing $\#(S(a,b) \cap S(b,c))$, we counted words with abc as a subword for $a,b,c \in A$ distinct). On the other hand, the intersection $P \cap Q$ is the set of words in S with 3701 as a substring. These can be built by choosing a character $a \in A$ (10 options) and a positition either before or after the string 3701 to place the character a (2 options). Hence $\#(P \cap Q) = 20$, and we use inclusion exclusion to compute

$$\#(X \cap Y \cap Z) = \#(P \cup Q) = \#P + \#Q - \#(P \cap Q) = 300 + 300 - 20 = 580.$$

Finally, putting this all together we have the following full count of S.

$$\begin{aligned} \#S = \#(X \cup Y \cup Z) &= \#X + \#Y + \#Z - \#(X \cap Y) - \#(X \cap Z) - \#(Y \cap Z) + \#(X \cap Y \cap Z) \\ &= 3070 + 3070 + 3070 - 357 - 60 - 357 + 20 = 8456 \end{aligned}$$

Problem 5. Let S be the set of length 5 words in alphabet $A = \{0\} \cup [9]$ with the subword 00. Determine $\#S$.

Problem 6. Let A be an alphabet of size k and $S_{n,k}$ the set of length n words in A with subword xx for a fixed $x \in A$. Count $\#S$.

Pigeonhole Principle

All of the above techniques for counting were used for exact counting. The following proposition, called the *Pigeonhole Principle*, is more an application of counting than a technique for counting.

Proposition 9 (Pigeonhole Principle)**.** *Let $f\colon A \to B$ be a function. If $\#A > \#B$, then there are $x, y \in A$ with $f(x) = f(y)$.*

Example 8. Suppose $S \subseteq \mathbb{Z}$ has cardinality $n \geq 2$. For all $m \in [n-2]$ there are distinct $x, y \in S$ such that $x \equiv y \pmod{m}$. To see this, let $f\colon S \to \{0\} \cup [m]$ be defined by $f(x)$ is the remainder of x modulo k; now $\#S = n > m+1 = \#(\{0\} \cup [m])$ yields that there are distinct $x, y \in S$ such that $f(x) = f(y)$ by the pigeonhole principle. Hence x and y have the same remainder modulo m yields $x \cong y \pmod{m}$ as desired.

The following is a nice generalization of the Pigeonhole Principle; its proof yields the Pigeonhole Principle as a corollary.

Proposition 10 (Generalized Pigeonhole Principle)**.** *Let $f\colon A \to B$ be a function with A and B finite sets, and let $k \in \mathbb{Z}_{>0}$. If $\#A > k\#B$, then there is a $b \in B$ with $\#f^{-1}(b) \geq k+1$.*

Proof. Let $f\colon A \to B$ be a function with A and B finite sets, and suppose there is $k \in \mathbb{Z}_{>0}$ such that $\#A > k\#B$. Assume to the contrary that $\#f^{-1}(b) \leq k$ for all $b \in B$. Note that $f^{-1}(b) \cap f^{-1}(b') = f^{-1}(\{b\} \cap \{b'\})$ by a previous result on function preimages. In particular, if $b \neq b'$ then $f^{-1}(b) \cap f^{-1}(b') = f^{-1}(\emptyset) = \emptyset$. Hence $\left\{f^{-1}(b) : b \in B\right\}$ is a collection pairwise disjoint subsets of $A = \operatorname{dom}(f)$. Moreover, as f is a function, for every $a \in A$ there is a unique $b \in B$ such that $f(a) = b$, i.e., $a \in f^{-1}(a)$. Thus $A = \bigcup_{b \in B} f^{-1}(b)$, and we may apply the Addition Principle our assumption above to compute as follows:

$$\#A = \#\bigcup_{b \in B} f^{-1}(b) = \sum_{b \in B} \#f^{-1}(b) \leq \sum_{b \in B} k = k \sum_{b \in B} 1 = k\#B.$$

This yields $k\#B < \#A \leq k\#B$, which is a contradiction.

We conclude there is some $b \in B$ for which $\#f^{-1}(b) \geq k+1$. □

Proof of Pigeonhole Principle. This follows from the Generalized Pigeonhole Principle with $k = 1$. □

Corollary 2. *For every function $f\colon A \to B$ for A and B finite sets, there exists a $b \in B$ such that $\#f^{-1}(b) \geq \left\lceil \frac{\#A}{\#B} \right\rceil$.*

Proof. Let $f\colon A \to B$ be a function of finite sets A and B. Note that $(\left\lceil \frac{x}{y} \right\rceil - 1)y < (\frac{x}{y} + 1 - 1)y = x$ for all $x, y \in \mathbb{R}_{>0}$ by basic properties of the ceiling function. Hence $(\left\lceil \frac{\#A}{\#B} \right\rceil - 1)\#B < \#A$, and thus there is a $b \in B$ such that $\#f^{-1}(b) \geq \left\lceil \frac{\#A}{\#B} \right\rceil$ by the Generalized Pigeonhole Principle. □

Problem 7. Let $S \subseteq \mathbb{Z}$ be arbitrary and finite with $\#S = n$.

1. Prove that for all $1 \leq d \leq 2^n - 2$ there are two nonempty subsets of S with the same sum modulo d.

Solution: Let $S \subseteq \mathbb{Z}$ be given and define $f\colon \mathbb{P}(S) \setminus \{\emptyset\} \to \mathbb{Z}/d\mathbb{Z}$ by $f(A) = \left[\sum_{a \in A} a\right]$ for all $\emptyset \neq A \subseteq S$. Note that $\#(\mathbb{Z}/d\mathbb{Z}) = d$ and $\#(\mathbb{P}(S) \setminus \{\emptyset\}) = 2^n - 1$. Hence there is a $k \in \mathbb{Z}/d\mathbb{Z}$ such that

$$\#f^{-1}(k) \geq \left\lceil \frac{2^n - 1}{d} \right\rceil \geq \left\lceil \frac{2^n - 1}{2^n - 2} \right\rceil = 2$$

by the Generalized Pigeonhole Principle. In particular, there are distinct nonempty subsets $A, B \subseteq S$ with $f(A) = f(B)$, and thus A and B belong to the same remainder class modulo d by definition of f.

NB: Our proof above actually does a little better than requested. We constructed at least $\left\lceil \frac{2^n - 1}{d} \right\rceil$ distinct nonempty subsets of S with the same sum modulo d.

2. Prove that for all $1 \leq d \leq n$ there is a nonempty subset $A \subseteq S$ with $\sum_{a \in A} a$ divisible by d.

Solution: Arbitrarily order the elements of S so that we may write $S = \{s_i : i \in [n]\}$. Define $S_k = \{s_i : i \in [k]\}$ for all $k \in [n]$; in particular, $S_1 = \{s_1\}$, $S_2 = \{s_1, s_2\}$, etc. Now let $1 \leq d \leq n-1$ and define $f\colon [n] \to \{0\} \cup [d-1]$ by $f(k) = r_k$ where r_k is the remainder of $\sum_{i=1}^{k} s_i$ under division by d for all $k \in [n]$. Note that if $f(k) = 0$ for some $k \in [n]$, then the sum of S_k is divisible by d immediately. Otherwise, by the Generalized Pigeonhole Principle there is $x \in \mathbb{Z}/d\mathbb{Z}$ such that $\#f^{-1}(x) \geq \left\lceil \frac{n}{d} \right\rceil \geq 2$. Thus there are distinct $1 \leq i < j \leq n$ with $f(i) = x = f(j)$; note that $S_i \subsetneq S_j$, so $S_j \setminus S_i \neq \emptyset$, and the remainders of $\sum_{k=1}^{i} s_k$ and $\sum_{k=1}^{j} s_k$ modulo d are equal by construction. Hence

$$\sum_{a \in S_j \setminus S_i} a = \sum_{k=i+1}^{j} s_k = \sum_{k=1}^{j} s_k - \sum_{k=1}^{i} s_k$$

is divisible by d by a previous proposition.

NB: The above argument can be strengthened for $1 \leq d \leq n-1$. We obtained distinct subsets of S with sum divisible by d for each pair $i \neq j$ with $f(i) = f(j)$; as $\#f^{-1}(x) \geq \left\lceil \frac{n}{d} \right\rceil$ we are guaranteed at least $\binom{\lceil \frac{n}{d} \rceil}{2}$ such pairs, and thus at least as many subsets $A \subseteq S$ with the desired property.

3. Prove that for all $n \in \mathbb{Z}_{>0}$ there is a positive integer with at most $n+1$ digits, all of which are either 0 or 1, which is divisible by n.

Solution: Letting $S = \left\{10^k : k \in [n]\right\}$, we appeal to the previous result with $d = n$. Thus there is a nonempty subset $A \subseteq S$ with $\sum_{a \in A} a$ divisible by d; but this number has digits only 0 and 1 and length at most $n+1$ by construction.

NB: This argument doesn't depend on the base-10 arithmetic system. Indeed, we could modify S to use powers of $b \in \mathbb{Z}_{\geq 2}$ to obtain the same result in arbitrary base-b arithmetic.

Problem 8. Let $n \in \mathbb{N}_0$ have $n \geq 3$. Find an upper bound on the minimum k such that every k-subset S of $[n]$ has two distinct subsets A and B with the same sum.

Solution: Note that the maximum sum of a k-subset $S \subseteq [n]$ is

$$\sum_{i=1}^{k} (n-i) = kn - \frac{k(k+1)}{2} = \frac{k(2n-k-1)}{2}.$$

Thus there are $\frac{k(2n-k-1)}{2}$ total possible sums for k-subsets of $[n]$. Note that $k \in [n]$ satisfies the above conditions provided $2^k = \#\mathbb{P}(S) > \frac{k(2n-k-1)}{2}$; indeed, if this equations is satisfied, then there are two subsets of S with the same sum by the Pigeonhole Principle. Rearranging the inequality above, we must have $2^{k+1} + k^2 > k(2n-1)$; thus $k_n = \min\left\{i \in \mathbb{Z}_{>0} : 2^{i+1} + i^2 > i(2n-1)\right\}$ is the general solution.

NB: The sequence k_n is small by comparison with n. Indeed, for $n = 10^{1000}$ we have $k_n = 3334$ and for $n = 10^{10000}$ we have $k_n = 33235$ (computed with a simple python program).